

Change Point in Panel Data with Small Fixed Panel Size: Ratio and Non-Ratio Test Statistics

Barbora Peřtová¹, Michal Peřta^{2*}

1. Institute of Computer Science, The Czech Academy of Sciences, Department of Medical Informatics and Biostatistics, Pod Vodárenskou věží 271/2, 18207 Prague, Czech Republic
2. Faculty of Mathematics and Physics, Charles University in Prague, Department of Probability and Mathematical Statistics, Sokolovská 49/83, 18675 Prague, Czech Republic

Abstract: The main goal is to develop and, consequently, compare stochastic methods for detection whether a structural change in panel data occurred at some unknown time or not. Panel data of our interest consist of a moderate or relatively large number of panels, while the panels contain a small number of observations. Testing procedures to detect a possible common change in means of the panels are established. Ratio and non-ratio type test statistics are considered. Their asymptotic distributions under the no change null hypothesis are derived. Moreover, we prove the consistency of the tests under the alternative. The main advantage of the ratio type statistics compared to the non-ratio ones is that the variance of the observations neither has to be known nor estimated. A simulation study reveals that the proposed ratio statistic outperforms the non-ratio one by keeping the significance level under the null, mainly when stronger dependence within the panel is taken into account. However, the non-ratio statistic rejects the null in the simulations more often than it should, which yields higher power compared to the ratio statistic.

Keywords: Change point, Panel data, Change in mean, Hypothesis testing, Structural change, Fixed panel size, Short panels, Ratio type statistics, Non-ratio type statistic, CUSUM type statistics.

1 Introduction

The problem of an unknown common change in means of the panels is studied here, where the panel data consist of N panels and each panel contains T observations over time. Various values of the change are possible for each panel at some unknown common time $\tau = 1, \dots, N$. The panels are considered to be independent, but this restriction can be weakened. In spite of that, observations within the panel are usually not independent. It is supposed that a common unknown dependence structure is present over the panels.

Tests for change point detection in the panel data have been proposed only in case when the panel size T is sufficiently large, i.e., T increases over all limits from an asymptotic point of view, cf. [Chan *et al.* \(2013\)](#) or [Horváth and Huřková \(2012\)](#). However, the change point estimation has already been studied for finite T not depending on the number of panels N , see [Bai \(2010\)](#) or [Peřtová and Peřta \(2016\)](#). The remaining task is to develop testing procedures to decide whether a common change point is present or not in the panels, while taking into account that the length T of each observation regime is fixed and can be relatively small.

2 Panel Change Point Model

Let us consider the panel change point model

$$Y_{i,t} = \mu_i + \delta_i \mathcal{I}\{t > \tau\} + \sigma \varepsilon_{i,t}, \quad 1 \leq i \leq N, 1 \leq t \leq T; \quad (1)$$

*Corresponding author: michal.pesta@mff.cuni.cz

where $\sigma > 0$ is an unknown variance-scaling parameter and T is fixed, not depending on N . The possible *common change point time* is denoted by $\tau \in \{1, \dots, T\}$. A situation where $\tau = T$ corresponds to *no change* in means of the panels. The means μ_i are panel-individual. The amount of the break in mean, which can also differ for every panel, is denoted by δ_i . Furthermore, it is assumed that the sequences of panel disturbances $\{\varepsilon_{i,t}\}_t$ are independent and within each panel the errors form a weakly stationary sequence with a common correlation structure. This can be formalized in the following assumption.

Assumption A1. The vectors $[\varepsilon_{i,1}, \dots, \varepsilon_{i,T}]^\top$ existing on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are *iid* for $i = 1, \dots, N$ with $\mathbb{E}\varepsilon_{i,t} = 0$ and $\text{Var } \varepsilon_{i,t} = 1$, having the autocorrelation function

$$\rho_t = \text{Corr}(\varepsilon_{i,s}, \varepsilon_{i,s+t}) = \text{Cov}(\varepsilon_{i,s}, \varepsilon_{i,s+t}), \quad \forall s \in \{1, \dots, T-t\},$$

which is independent of the lag s , the cumulative autocorrelation function

$$r(t) = \text{Var} \sum_{s=1}^t \varepsilon_{i,s} = \sum_{|s| < t} (t - |s|) \rho_s,$$

and the shifted cumulative correlation function

$$R(t, v) = \text{Cov} \left(\sum_{s=1}^t \varepsilon_{i,s}, \sum_{u=t+1}^v \varepsilon_{i,u} \right) = \sum_{s=1}^t \sum_{u=t+1}^v \rho_{u-s}, \quad t < v$$

for all $i = 1, \dots, N$ and $t, v = 1, \dots, T$.

The sequence $\{\varepsilon_{i,t}\}_{t=1}^T$ can be viewed as a part of a *weakly stationary* process. Note that the dependent errors within each panel do not necessarily need to be linear processes. For example, GARCH processes as error sequences are allowed as well. The assumption of independent panels can indeed be relaxed, but it would make the setup much more complex. Consequently, probabilistic tools for dependent data need to be used (e.g., suitable versions of the central limit theorem). Nevertheless, assuming, that the claim amounts for different insurance companies are independent, is reasonable. Moreover, the assumption of a common homoscedastic variance parameter σ can be generalized by introducing weights $w_{i,t}$, which are supposed to be known. Being particular in actuarial practice, it would mean to normalize the total claim amount by the premium received, since bigger insurance companies are expected to have higher variability in total claim amounts paid.

It is required to test the *null hypothesis* of no change in the means

$$H_0 : \tau = T$$

against the *alternative* that at least one panel has a change in mean

$$H_1 : \tau < T \quad \text{and} \quad \exists i \in \{1, \dots, N\} : \delta_i \neq 0.$$

3 Test Statistic and Asymptotic Results

We propose a *ratio type statistic* to test H_0 against H_1 , because this type of statistic does not require estimation of the nuisance parameter for the variance. Generally, this is due to the fact that the variance parameter simply cancels out from the nominator and denominator of the statistic. For surveys on ratio type test statistics, we refer to [Csörgő and Horváth \(1997\)](#),

Madurkayová (2011), and Horváth *et al.* (2009). Our particular panel change point CUSUM test statistic is

$$\mathcal{C}_N(T) = \frac{1}{\sqrt{N}} \max_{t=1, \dots, T-1} \left| \sum_{i=1}^N \sum_{s=1}^t (Y_{i,s} - \bar{Y}_{i,t}) \right|,$$

which is going to be compared with the ratio test statistics

$$\mathcal{R}_N(T) = \max_{t=2, \dots, T-2} \frac{\max_{s=1, \dots, t} \left| \sum_{i=1}^N \sum_{r=1}^s (Y_{i,r} - \bar{Y}_{i,t}) \right|}{\max_{s=t, \dots, T-1} \left| \sum_{i=1}^N \sum_{r=s+1}^T (Y_{i,r} - \tilde{Y}_{i,t}) \right|},$$

where $\bar{Y}_{i,t}$ is the average of the first t observations in panel i and $\tilde{Y}_{i,t}$ is the average of the last $T - t$ observations in panel i , i.e.,

$$\bar{Y}_{i,t} = \frac{1}{t} \sum_{s=1}^t Y_{i,s} \quad \text{and} \quad \tilde{Y}_{i,t} = \frac{1}{T-t} \sum_{s=t+1}^T Y_{i,s}.$$

elaborated in Peřtová and Peřta (2015). It will be demonstrated by simulations that $\mathcal{R}_N(T)$ keeps the theoretical significance level, however, $\mathcal{C}_N(T)$ does not.

Firstly, we derive the behavior of the test statistics under the null hypothesis.

Theorem 1 (Under Null). *Under hypothesis H_0 and Assumption A1*

$$\mathcal{C}_N(T) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \sigma \max_{t=1, \dots, T-1} \left| X_t - \frac{t}{T} X_T \right|$$

and

$$\mathcal{R}_N(T) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \max_{t=2, \dots, T-2} \frac{\max_{s=1, \dots, t} \left| X_s - \frac{s}{t} X_t \right|}{\max_{s=t, \dots, T-1} \left| Z_s - \frac{T-s}{T-t} Z_t \right|},$$

where $Z_t := X_T - X_t$ and $[X_1, \dots, X_T]^\top$ is a multivariate normal random vector with zero mean and covariance matrix $\mathbf{\Lambda} = \{\lambda_{t,v}\}_{t,v=1}^{T,T}$ such that

$$\lambda_{t,t} = r(t) \quad \text{and} \quad \lambda_{t,v} = r(t) + R(t, v), \quad t < v.$$

The limiting distribution does not depend on the variance nuisance parameter σ , but it depends on the unknown correlation structure of the panel change point model, which has to be estimated for testing purposes. The way of its estimation is shown in Section 4. Note that in case of independent observations within the panel, the correlation structure and, hence, the covariance matrix $\mathbf{\Lambda}$ is simplified such that $r(t) = t$ and $R(t, v) = 0$.

Next, we show how the test statistic behaves under the alternative.

Assumption A2. $\lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \left| \sum_{i=1}^N \delta_i \right| = \infty$.

Theorem 2 (Under Alternative). *If $\tau \leq T - 3$, then under Assumptions A1, A2 and alternative H_1*

$$\mathcal{C}_N(T) \xrightarrow[N \rightarrow \infty]{P} \infty \quad \text{and} \quad \mathcal{R}_N(T) \xrightarrow[N \rightarrow \infty]{P} \infty. \quad (2)$$

Assumption A2 is satisfied, for instance, if $0 < \delta \leq \delta_i \forall i$ (a common lower change point threshold) and $\delta \sqrt{N} \rightarrow \infty$, $N \rightarrow \infty$. Another suitable example of δ_i s for the condition in Assumption A2, can be $0 < \delta_i = KN^{-1/2+\eta}$ for some $K > 0$ and $\eta > 0$. Or $\delta_i = Ci^{\alpha-1}\sqrt{N}$ may be used as well, where $\alpha \geq 0$ and $C > 0$. The assumption $\tau \leq T - 3$ means that there

are at least three observations in the panel after the change point. It is also possible to redefine the test statistic by interchanging the nominator and the denominator of $\mathcal{S}_N(T)$. Afterwards, Theorem 2 for the modified test statistic would require three observations before the change point, i.e., $\tau \geq 3$.

Theorem 2 says that in presence of a structural change in the panel means, the test statistic explodes above all bounds. Hence, the procedure is consistent and the asymptotic distribution from Theorem 1 can be used to construct the test.

4 Estimation of the Covariance Structure

The estimation of the covariance matrix $\mathbf{\Lambda}$ from Theorem 1 requires panels as vectors with elements having common mean (i.e., without a jump). Therefore, it is necessary to construct an estimate for a possible change point. A *consistent estimate* of the change point τ in the panel data is proposed in Peřtová and Peřta (2016) as

$$\hat{\tau}_N := \arg \min_{t=2, \dots, T} \frac{1}{w(t)} \sum_{i=1}^N \sum_{s=1}^t (Y_{i,s} - \bar{Y}_{i,t})^2, \quad (3)$$

where $\{w(t)\}_{t=2}^T$ is a sequence of weights specified in Peřtová and Peřta (2016).

Since the panels are considered to be independent and the number of panels may be sufficiently large, one can estimate the correlation structure of the errors $[\varepsilon_{1,1}, \dots, \varepsilon_{1,T}]^\top$ empirically. We base the errors' estimates on *residuals*

$$\hat{e}_{i,t} := \begin{cases} Y_{i,t} - \bar{Y}_{i,\hat{\tau}_N}, & t \leq \hat{\tau}_N, \\ Y_{i,t} - \tilde{Y}_{i,\hat{\tau}_N}, & t > \hat{\tau}_N. \end{cases} \quad (4)$$

Then, the empirical version of the autocorrelation function is

$$\hat{\rho}_t := \frac{1}{\hat{\sigma}^2 NT} \sum_{i=1}^N \sum_{s=1}^{T-t} \hat{e}_{i,s} \hat{e}_{i,s+t}.$$

Consequently, the kernel estimation of the cumulative autocorrelation function and shifted cumulative correlation function is adopted in lines with Andrews (1991):

$$\begin{aligned} \hat{r}(t) &= \sum_{|s| < t} (t - |s|) \kappa\left(\frac{s}{h}\right) \hat{\rho}_s, \\ \hat{R}(t, v) &= \sum_{s=1}^t \sum_{u=t+1}^v \kappa\left(\frac{u-s}{h}\right) \hat{\rho}_{u-s}, \quad t < v; \end{aligned}$$

where $h > 0$ stands for the window size and κ belongs to a class of kernels

$$\left\{ \kappa(\cdot) : \mathbb{R} \rightarrow [-1, 1] \mid \kappa(0) = 1, \kappa(x) = \kappa(-x), \forall x, \int_{-\infty}^{+\infty} \kappa^2(x) dx < \infty, \right. \\ \left. \kappa(\cdot) \text{ is continuous at } 0 \text{ and at all but a finite number of other points} \right\}.$$

Since the variance parameter σ simply cancels out from the limiting distribution of Theorem 1, it neither has to be estimated nor known. Nevertheless, one can use $\hat{\sigma}^2 := \frac{1}{NT} \sum_{i=1}^N \sum_{s=1}^T \hat{e}_{i,s}^2$.

5 Simulations

A simulation experiment was performed to study the *finite sample* properties of the test statistics for a common change in panel means. In particular, the interest lies in the empirical *sizes* of the proposed tests (i.e., based on $\mathcal{R}_N(T)$ and $\mathcal{S}_N(T)$) under the null hypothesis and in the empirical *rejection* rate (power) under the alternative. Random samples of panel data (5000 each time) are generated from the panel change point model (1). The panel size is set to $T = 10$ and $T = 25$ in order to demonstrate the performance of the testing approaches in case of small and intermediate panel length. The number of panels considered is $N = 50$ and $N = 200$.

The correlation structure within each panel is modeled via random vectors generated from iid, AR(1), and GARCH(1,1) sequences. The considered AR(1) process has coefficient $\phi = 0.3$. In case of GARCH(1,1) process, we use coefficients $\alpha_0 = 1$, $\alpha_1 = 0.1$, and $\beta_1 = 0.2$, which according to Lindner (2009, Example 1) gives a strictly stationary process. In all three sequences, the innovations are obtained as iid random variables from a standard normal $N(0, 1)$ or Student t_5 distribution. Simulation scenarios are produced as all possible combinations of the above mentioned settings.

When using the asymptotic distribution from Theorem 1, the covariance matrix is estimated as proposed in Section 4 using the Parzen kernel

$$\kappa_P(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, & 0 \leq |x| \leq 1/2; \\ 2(1 - |x|)^3, & 1/2 \leq |x| \leq 1; \\ 0, & \text{otherwise.} \end{cases}$$

Bartlett window as well

Several values of the smoothing window width h are tried from the interval $[2, 5]$ and all of them work fine providing comparable results. To simulate the asymptotic distribution of the test statistics, 2000 multivariate random vectors are generated using the pre-estimated covariance matrix. To access the theoretical results under H_0 numerically, Table 1 provides the empirical specificity (one minus size) of the asymptotic tests based on $\mathcal{R}_N(T)$ and $\mathcal{S}_N(T)$, where the significance level is $\alpha = 5\%$.

Table 1: Empirical specificity (1–size) of the test under H_0 for test statistics $\mathcal{R}_N(T)$ and $\mathcal{C}_N(T)$ using the asymptotic critical values, considering a significance level of 5%, weight function $w(t) = t^2$, and smoothing window width $h = 2$

T	N	innovations	IID		AR(1)		GARCH(1,1)	
10	50	$N(0, 1)$.948	.934	.933	.823	.946	.935
		t_5	.951	.925	.932	.822	.946	.929
	200	$N(0, 1)$.950	.933	.939	.825	.950	.938
		t_5	.948	.927	.935	.821	.948	.937
25	50	$N(0, 1)$.946	.940	.932	.780	.947	.945
		t_5	.948	.945	.932	.790	.946	.943
	200	$N(0, 1)$.949	.939	.930	.801	.951	.939
		t_5	.953	.941	.931	.813	.952	.946

It may be seen that both approaches are close to the theoretical value of specificity .95. As expected, the best results are achieved in case of independence within the panel, because there

is no information overlap between two consecutive observations. The precision of not rejecting the null is increasing as the number of panels is getting higher and the panel is getting longer as well.

The performance of both testing procedures under H_1 in terms of the empirical rejection rates is shown in Table 2, where the change point is set to $\tau = \lfloor T/2 \rfloor$ and the change sizes δ_i are independently uniform on $[1, 3]$ in 33%, 66% or in all panels.

Table 2: Empirical sensitivity (power) of the test under H_1 for test statistics $\mathcal{R}_N(T)$ and $\mathcal{C}_N(T)$ using the asymptotic critical values, considering a significance level of 5%, weight function $w(t) = t^2$, and smoothing window width $h = 2$

H_1	T	N	innovations	IID		AR(1)		GARCH(1,1)	
33%	10	50	$N(0, 1)$.235	1.000	.256	.999	.193	1.000
			t_5	.174	.999	.202	.996	.201	.999
		200	$N(0, 1)$.453	1.000	.486	1.000	.387	1.000
			t_5	.360	1.000	.393	1.000	.389	1.000
		25	$N(0, 1)$.376	1.000	.394	.992	.312	1.000
			t_5	.294	1.000	.301	.993	.312	1.000
	25	50	$N(0, 1)$.685	1.000	.699	.995	.584	1.000
			t_5	.561	1.000	.565	1.000	.590	1.000
		200	$N(0, 1)$.774	1.000	.807	1.000	.677	1.000
			t_5	.642	1.000	.692	1.000	.688	1.000
		25	$N(0, 1)$.688	1.000	.694	1.000	.581	1.000
			t_5	.558	1.000	.570	1.000	.594	1.000
66%	10	50	$N(0, 1)$.450	1.000	.491	1.000	.386	1.000
			t_5	.360	1.000	.377	1.000	.390	1.000
		200	$N(0, 1)$.774	1.000	.807	1.000	.677	1.000
			t_5	.642	1.000	.692	1.000	.688	1.000
		25	$N(0, 1)$.688	1.000	.694	1.000	.581	1.000
			t_5	.558	1.000	.570	1.000	.594	1.000
	25	50	$N(0, 1)$.951	1.000	.959	1.000	.905	1.000
			t_5	.874	1.000	.888	1.000	.906	1.000
		200	$N(0, 1)$.928	1.000	.945	1.000	.868	1.000
			t_5	.844	1.000	.869	1.000	.872	1.000
		25	$N(0, 1)$.873	1.000	.884	1.000	.792	1.000
			t_5	.760	1.000	.771	1.000	.789	1.000
100%	10	50	$N(0, 1)$.997	1.000	.997	1.000	.985	1.000
			t_5	.977	1.000	.982	1.000	.986	1.000
		200	$N(0, 1)$.997	1.000	.997	1.000	.985	1.000
			t_5	.977	1.000	.982	1.000	.986	1.000
		25	$N(0, 1)$.997	1.000	.997	1.000	.985	1.000
			t_5	.977	1.000	.982	1.000	.986	1.000

One can conclude that the power of both tests increases as the panel size and the number of panels increase, which is straightforward and expected. Moreover, higher power is obtained when a larger portion of panels is subject to have a change in mean. The test power drops when switching from independent observations within the panel to dependent ones. Innovations with heavier tails (i.e., t_5) yield smaller power than innovations with lighter tails. Generally, the newly defined test statistic $\mathcal{S}_N(T)$ outperforms $\mathcal{R}_N(T)$ in all scenarios with respect to the power.

Finally, an early change point is discussed very briefly. We stay with standard normal innovations, iid observations within the panel, the size of changes δ_i being independently uniform on

[1, 3] in all panels, and the change point is $\tau = 3$ in case of $T = 10$ and $\tau = 5$ for $T = 25$. The empirical sensitivities of both tests for small values of τ are shown in Table 3.

Table 3: Empirical sensitivity of the test for small values of τ under H_1 for test statistics $\mathcal{R}_N(T)$ and $\mathcal{C}_N(T)$ using the asymptotic critical values, considering a significance level of 5%, weight function $w(t) = t^2$, and smoothing window width $h = 2$

T	τ	N	H_1 , iid, $N(0, 1)$		T	τ	N	H_1 , iid, $N(0, 1)$	
10	3	50	.551	1.000	25	5	50	.629	1.000
		200	.867	1.000			200	.927	1.000

When the change point is not in the middle of the panel, the power of the test generally falls down. The source of such decrease is that the left or right part of the panel possesses less observations with constant mean, which leads to a decrease of precision in the correlation estimation. Nevertheless, $\mathcal{S}_N(T)$ again outperforms $\mathcal{R}_N(T)$ even for early or late change points (the late change points are not numerically demonstrated here).

6 Conclusions

In this paper, we consider the change point problem in panel data with fixed panel size. Occurrence of common breaks in panel means is tested. We introduce a ratio type test statistic and derive its asymptotic properties. Under the null hypothesis of no change, the test statistic weakly converges to a functional of the multivariate normal random vector with zero mean and covariance structure depending on the intra-panel covariances. As shown in the paper, these covariances can be estimated and, consequently, used for testing whether a change in means occurred or not. This is indeed feasible, because the test statistic under the alternative converges to infinity in probability. Furthermore, the whole stochastic theory behind requires relatively simple assumptions, which are not too restrictive.

A simulation study illustrates that even for small panel size, both investigated approaches—the newly derived one based on $\mathcal{S}_N(T)$ and the older one proposed in [Peřtová and Peřta \(2015\)](#)—work fine. One may judge that both methods keep the significance level under the null, while various simulation scenarios are considered. Besides that, the power of the test is higher in case of $\mathcal{S}_N(T)$ compared to $\mathcal{R}_N(T)$. Finally, the proposed method is applied to insurance data, for which the change point analysis in panel data provides an appealing approach.

6.1 Discussion

Our setup can be modified by considering large panel size, i.e., $T \rightarrow \infty$. Consequently, the whole theory leads to convergences to functionals of Gaussian processes with a covariance structure derived in a very similar fashion as for fixed T . However, our motivation is to develop techniques for fixed and relatively small panel size.

Dependent panels may be taken into account and the presented work might be generalized for some kind of asymptotic independence over the panels or prescribed dependence among the panels. Nevertheless, our incentive is determined by a problem from non-life insurance, where the association of insurance companies consists of a relatively high number of insurance companies. Thus, the portfolio of yearly claims is so diversified, that the panels corresponding to

insurance companies' yearly claims may be viewed as independent and neither natural ordering nor clustering has to be assumed.

A Proofs

Proof of Theorem 1. Let us define

$$U_N(t) := \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N \sum_{s=1}^t (Y_{i,s} - \mu_i).$$

Using the multivariate Lindeberg-Lévy CLT for a sequence of T -dimensional iid random vectors $\{[\sum_{s=1}^1 \varepsilon_{i,s}, \dots, \sum_{s=1}^T \varepsilon_{i,s}]^\top\}_{i \in \mathbb{N}}$, we have under H_0

$$[U_N(1), \dots, U_N(T)]^\top \xrightarrow[N \rightarrow \infty]{\mathcal{D}} [X_1, \dots, X_T]^\top,$$

since $\text{Var} [\sum_{s=1}^1 \varepsilon_{1,s}, \dots, \sum_{s=1}^T \varepsilon_{1,s}]^\top = \mathbf{\Lambda}$. Indeed, the t -th diagonal element of the covariance matrix $\mathbf{\Lambda}$ is $\text{Var} \sum_{s=1}^t \varepsilon_{1,s} = r(t)$ and the upper off-diagonal element on position (t, v) is

$$\begin{aligned} \text{Cov} \left(\sum_{s=1}^t \varepsilon_{1,s}, \sum_{u=1}^v \varepsilon_{1,u} \right) &= \text{Var} \sum_{s=1}^t \varepsilon_{1,s} + \text{Cov} \left(\sum_{s=1}^t \varepsilon_{1,s}, \sum_{u=t+1}^v \varepsilon_{1,u} \right) \\ &= r(t) + R(t, v), \quad t < v. \end{aligned}$$

Moreover, let us define the reverse analogue to $U_N(t)$, i.e.,

$$V_N(t) := \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N \sum_{s=t+1}^T (Y_{i,s} - \mu_i) = U_N(T) - U_N(t).$$

Hence,

$$\begin{aligned} U_N(s) - \frac{s}{t} U_N(t) &= \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N \left\{ \sum_{r=1}^s \left[(Y_{i,r} - \mu_i) - \frac{1}{t} \sum_{v=1}^t (Y_{i,v} - \mu_i) \right] \right\} \\ &= \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N \sum_{r=1}^s (Y_{i,r} - \bar{Y}_{i,t}) \end{aligned}$$

and, consequently,

$$\begin{aligned} V_N(s) - \frac{T-s}{T-t} V_N(t) &= \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N \left\{ \sum_{r=s+1}^T \left[(Y_{i,r} - \mu_i) - \frac{1}{T-t} \sum_{v=t+1}^T (Y_{i,v} - \mu_i) \right] \right\} \\ &= \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N \sum_{r=s+1}^T (Y_{i,r} - \tilde{Y}_{i,t}). \end{aligned}$$

Using the Cramér-Wold device, we end up with

$$\max_{t=1, \dots, T-1} \left| U_N(t) - \frac{t}{T} U_N(T) \right| \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \max_{t=1, \dots, T-1} \left| X_t - \frac{t}{T} X_T \right|$$

and

$$\begin{aligned} \max_{t=2,\dots,T-2} \frac{\max_{s=1,\dots,t} |U_N(s) - \frac{s}{t} U_N(t)|}{\max_{s=t,\dots,T-1} |V_N(s) - \frac{T-s}{T-t} V_N(t)|} \\ \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \max_{t=2,\dots,T-2} \frac{\max_{s=1,\dots,t} |X_s - \frac{s}{t} X_t|}{\max_{s=t,\dots,T-1} |(X_T - X_s) - \frac{T-s}{T-t} (X_T - X_t)|}. \end{aligned}$$

□

Proof of Theorem 2. Considering $\mathcal{C}_N(T)$, we have under alternative H_1 that

$$\begin{aligned} \frac{1}{\sigma\sqrt{N}} \left| \sum_{i=1}^N \sum_{s=1}^{\tau} (Y_{i,s} - \bar{Y}_{i,T}) \right| &= \frac{1}{\sigma\sqrt{N}} \left| \sum_{i=1}^N \sum_{s=1}^{\tau} \left(\mu_i + \sigma\varepsilon_{i,s} - \frac{1}{T} \sum_{v=1}^T (\mu_i + \sigma\varepsilon_{i,v}) - \frac{1}{T} \delta_i \right) \right| \\ &= \frac{1}{\sqrt{N}} \left| \sum_{i=1}^N \sum_{s=1}^{\tau} (\varepsilon_{i,s} - \bar{\varepsilon}_{i,T}) - \frac{\tau}{\sigma T} \sum_{i=1}^N \delta_i \right| = \mathcal{O}_P(1) + \frac{\tau}{\sigma T\sqrt{N}} \left| \sum_{i=1}^N \delta_i \right| \xrightarrow{P} \infty, \quad N \rightarrow \infty, \end{aligned}$$

where $\bar{\varepsilon}_{i,T} = \frac{1}{\tau} \sum_{v=1}^T \varepsilon_{i,v}$.

In case of $\mathcal{R}_N(T)$, let $t = \tau + 1$. Then under alternative H_1 , it holds that

$$\begin{aligned} \frac{1}{\sigma\sqrt{N}} \max_{s=1,\dots,\tau+1} \left| \sum_{i=1}^N \sum_{r=1}^s (Y_{i,r} - \bar{Y}_{i,\tau+1}) \right| &\geq \frac{1}{\sigma\sqrt{N}} \left| \sum_{i=1}^N \sum_{r=1}^{\tau} (Y_{i,r} - \bar{Y}_{i,\tau+1}) \right| \\ &= \frac{1}{\sigma\sqrt{N}} \left| \sum_{i=1}^N \sum_{r=1}^{\tau} \left(\mu_i + \sigma\varepsilon_{i,r} - \frac{1}{\tau+1} \sum_{v=1}^{\tau+1} (\mu_i + \sigma\varepsilon_{i,v}) - \frac{1}{\tau+1} \delta_i \right) \right| \\ &= \frac{1}{\sqrt{N}} \left| \sum_{i=1}^N \sum_{r=1}^{\tau} (\varepsilon_{i,r} - \bar{\varepsilon}_{i,\tau+1}) - \frac{\tau}{\sigma(\tau+1)} \sum_{i=1}^N \delta_i \right| = \mathcal{O}_P(1) + \frac{\tau}{\sigma(\tau+1)\sqrt{N}} \left| \sum_{i=1}^N \delta_i \right| \xrightarrow{P} \infty, \quad N \rightarrow \infty, \end{aligned}$$

where $\bar{\varepsilon}_{i,\tau+1} = \frac{1}{\tau+1} \sum_{v=1}^{\tau+1} \varepsilon_{i,v}$.

Since there is no change after $\tau + 1$ and $\tau \leq T - 3$, then by Theorem 1 we obtain

$$\frac{1}{\sigma\sqrt{N}} \max_{s=\tau+1,\dots,T-1} \left| \sum_{i=1}^N \sum_{r=s+1}^T (Y_{i,r} - \tilde{Y}_{i,\tau+1}) \right| \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \max_{s=\tau+1,\dots,T-1} \left| Z_s - \frac{T-s}{T-\tau} Z_{\tau+1} \right|.$$

□

Acknowledgment

Institutional support to Barbora Peřtová was provided by RVO:67985807. The research of Michal Peřta was supported by the Czech Science Foundation project “DYME – Dynamic Models in Economics” No. P402/12/G097.

References

- Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59**(3), 817–858.
- Bai, J. (2010). Common breaks in means and variances for panel data. *Journal of Econometrics*, **157**(1), 78–92.

- Chan, J., Horváth, L., and Hušková, M. (2013). Change-point detection in panel data. *Journal of Statistical Planning and Inference*, **143**(5), 955–970.
- Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. Wiley, Chichester.
- Horváth, L. and Hušková, M. (2012). Change-point detection in panel data. *Journal of Time Series Analysis*, **33**(4), 631–648.
- Horváth, L., Horváth, Z., and Hušková, M. (2009). Ratio tests for change point detection. In N. Balakrishnan, E. A. Peña, and M. J. Silvapulle, editors, *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen*, volume 1, pages 293–304, Beachwood, Ohio. IMS Collections.
- Lindner, A. M. (2009). Stationarity, mixing, distributional properties and moments of GARCH(p,q)-processes. In T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch, editors, *Handbook of Financial Time Series*, pages 481–496, Berlin. Springer.
- Madurkayová, B. (2011). Ratio type statistics for detection of changes in mean. *Acta Universitatis Carolinae: Mathematica et Physica*, **52**(1), 47–58.
- Peštová, B. and Pešta, M. (2015). Testing structural changes in panel data with small fixed panel size and bootstrap. *Metrika*, **78**(6), 665–689.
- Peštová, B. and Pešta, M. (2016). Erratum to: Testing structural changes in panel data with small fixed panel size and bootstrap. *Metrika*, **79**(2), 237–238.